



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

Στατιστική στην Εκπαίδευση II

Σημειώσεις «Εισαγωγή στο SPSS» (δημιουργία βάσεων δεδομένων, πράξεις μεταβλητών)

Διδάσκων: Μιχάλης Λιναρδάκης
ΠΑΙΔΑΓΩΓΙΚΟ ΤΜΗΜΑ ΔΗΜΟΤΙΚΗΣ ΕΚΠΑΙΔΕΥΣΗΣ
ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΑΓΩΓΗΣ



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται στην άδεια χρήσης **Creative Commons** και ειδικότερα ***Αναφορά – Μη εμπορική Χρήση – Όχι Παράγωγο Έργο 3.0 Ελλάδα*** (***Attribution – Non Commercial – Non-derivatives 3.0 Greece***)



[ή επιλογή ενός άλλου από τους έξι συνδυασμούς]

[και αντικατάσταση λογότυπου άδειας όπου αυτό έχει μπει (σελ. 1, σελ. 2 και τελευταία)]

- Εξαιρείται από την ως άνω άδεια υλικό που περιλαμβάνεται στις διαφάνειες του μαθήματος, και υπόκειται σε άλλου τύπου άδεια χρήσης. Η άδεια χρήσης στην οποία υπόκειται το υλικό αυτό αναφέρεται ρητώς.

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Κρήτης**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

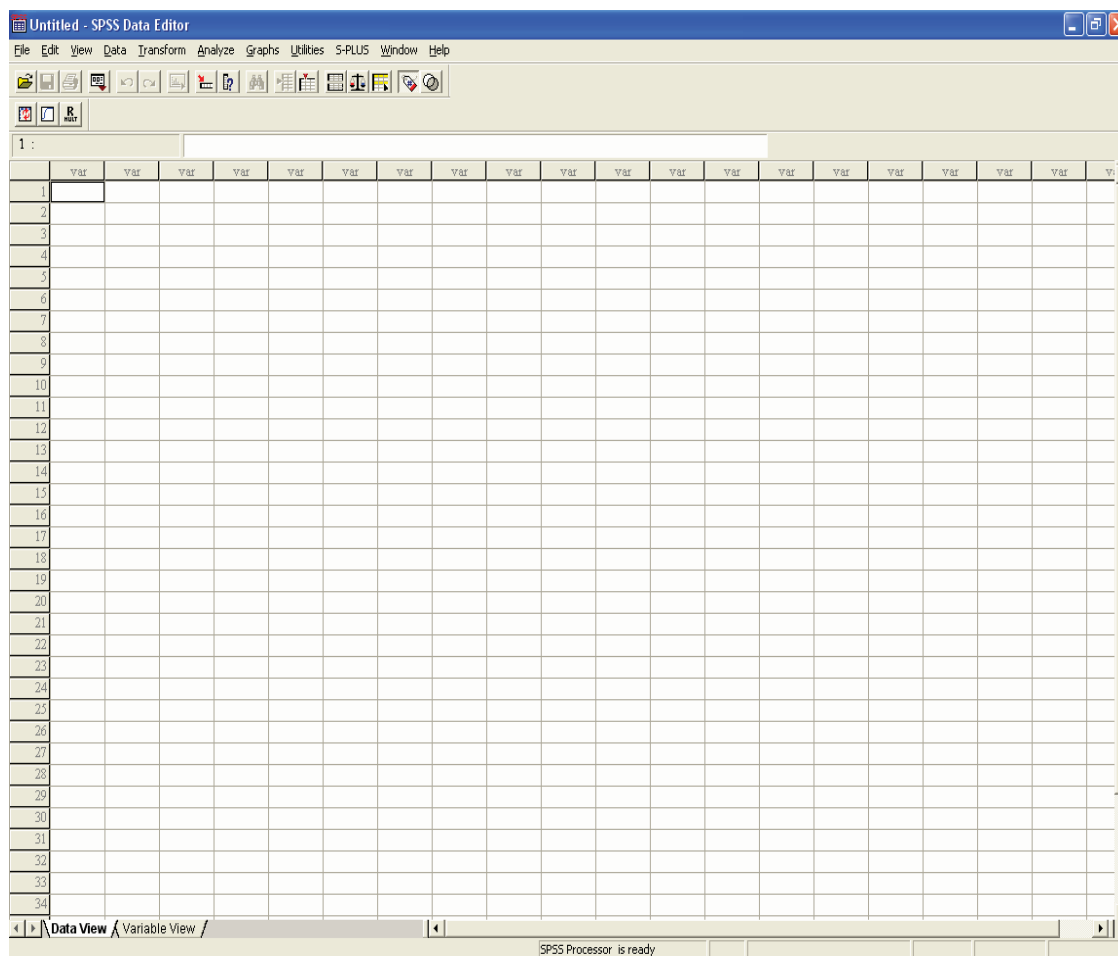


ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Παράρτημα 1 Εισαγωγή στο SPSS

Εισαγωγή δεδομένων

Η αρχική οθόνη του SPSS προβάλλει τον πίνακα όπου θα εισαχθούν τα δεδομένα. Εξ ορισμού, κατά την εκκίνηση εμφανίζεται το παράθυρο “Data view”. Κάθε γραμμή αντιστοιχεί σε μια παρατήρηση του δείγματος και κάθε στήλη σε μια μεταβλητή. Μετά την εισαγωγή τους, τα αριθμητικά δεδομένα στοιχίζονται δεξιά ενώ τα αλφαριθμητικά αριστερά. Αυτό είναι σημαντικό επειδή αν σε μια αριθμητική μεταβλητή, εισαχθεί κατά λάθος ένας χαρακτήρας (για παράδειγμα, αν αντί για το σύμβολο «,» ως υποδιαστολή σε εισαγωγή αριθμών με δεκαδικά, εισαχθεί το σύμβολο «.»), η μεταβλητή μετατρέπεται αυτόματα σε αλφαριθμητική (στοιχίζεται στο αριστερό μέρος της στήλης) και, κατά συνέπεια, δεν θα είναι δυνατή η περαιτέρω χρησιμοποίησή της σε στατιστικές αναλύσεις.



Έστω ότι έχουμε τέσσερις μεταβλητές και πέντε άτομα σε ένα σετ δεδομένων. Οι μεταβλητές που περιλαμβάνονται είναι οι

- Βάρος (σε κιλά, συνεχής μεταβλητή)
- Ύψος (σε εκατοστά, συνεχής μεταβλητή)
- Φύλο (με τιμές 1 και 2 για άνδρες και γυναίκες αντίστοιχα, κατηγορική-ονομαστική μεταβλητή)
- αριθμός ταυτότητας (αλφαριθμητική μεταβλητή).

Με την εισαγωγή των δεδομένων στο SPSS θα προκύψει το επόμενο παράθυρο:

	VAR00001	VAR00002	VAR00003	VAR00004
1	60.00	163.00	2.00	M111222
2	85.00	185.00	1.00	E333444
3	70.00	178.00	1.00	K444555
4	81.00	175.00	1.00	A222333
5	54.00	157.00	2.00	M444222

Επιλέγοντας “Variable view” στο αριστερό κάτω μέρος του παραθύρου, δίδεται το παράθυρο που ακολουθεί, το οποίο επιτρέπει την επεξεργασία των μεταβλητών.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	VAR00001	Numeric	8	2		None	None	10	Right	Scale
2	VAR00002	Numeric	8	2		None	None	10	Right	Scale
3	VAR00003	Numeric	8	2		None	None	9	Right	Scale
4	VAR00004	String	7	0		None	None	11	Left	Nominal

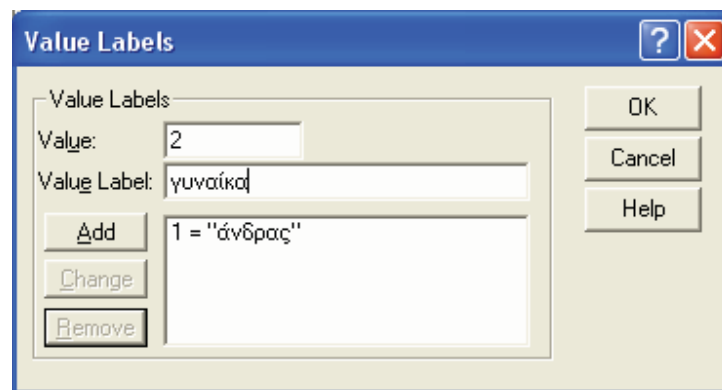
Σε όλες τις στήλες μπορούμε να εισάγουμε και ελληνικούς χαρακτήρες, αρκεί να επιλέξουμε τις στήλες αυτές και στη συνέχεια, από το menu “View” και “Fonts” να εισάγουμε μια γραμματοσειρά που υποστηρίζει ελληνικούς χαρακτήρες (π.χ. Times New Roman, Arial κ.λπ.).

Αναλυτικά στις στήλες:

- “Name”: εισάγεται το όνομα της μεταβλητής. Δίδεται ένα σύντομο όνομα που περιγράφει τη μεταβλητή. Δεν επιτρέπονται ειδικοί χαρακτήρες όπως το κόμμα, το ερωτηματικό, το κενό κ.λπ.
- “Type”: εισάγεται ο τύπος της μεταβλητής. Δηλώνονται ως “Numeric” οι αριθμητικές μεταβλητές (συνεχείς, κατηγορικές, μεταβλητές διάταξης) και ως “String” οι αλφαριθμητικές μεταβλητές. Επίσης, παρέχεται η δυνατότητα να δηλωθούν μεταβλητές που δηλώνουν ημερομηνία, χρηματικά ποσά με συμβολισμό νομίσματος κ.λπ.

- ➔ “Width”: δηλώνονται οι χαρακτήρες που καταλαμβάνει η μεταβλητή στο αρχείο των δεδομένων.
- ➔ “Decimals”: δηλώνεται ο αριθμός των δεκαδικών που θα εμφανίζονται στη μεταβλητή. Το πακέτο εξ ορισμού εμφανίζει 2 δεκαδικά σε όλες τις αριθμητικές μεταβλητές.
- ➔ “Label”: δηλώνεται ο τίτλος της κάθε μεταβλητής, ακόμη και επεξηγηματικά σχόλια. Μπορεί να περιλαμβάνει και ειδικούς χαρακτήρες.
- ➔ “Values”: χρησιμοποιείται στις κατηγορικές μεταβλητές για να δηλωθεί τι σημαίνουν οι κωδικοί που χρησιμοποιούνται (στο παράδειγμα μας, το «1» σημαίνει «άνδρας» και το «2» σημαίνει «γυναίκα»).
- ➔ “Missing”: κατά την εισαγωγή των δεδομένων, οι τιμές που λείπουν, δηλαδή που δεν έχουν καταγραφεί (missing values) μπορούν να δηλώνονται με αντίστοιχο κενό κελί (δηλαδή να μην εισάγεται κανένας χαρακτήρας σε αυτό). Αν παρόλα αυτά, θέλουμε να εισάγουμε ένα κωδικό ο οποίος θα υποδηλώνει τις τιμές που λείπουν, όπως για παράδειγμα τον κωδικό 999 που θα σημαίνει ότι η τιμή δεν έχει καταγραφεί, τότε αυτός ο κωδικός θα πρέπει να δηλωθεί στη στήλη “missing”.
- ➔ “Columns”: δηλώνεται το πλάτος που καταλαμβάνει η μεταβλητή στο αρχείο των δεδομένων.
- ➔ “Align”: δηλώνεται αν οι τιμές της μεταβλητής στοιχίζονται αριστερά, στο κέντρο ή δεξιά. Εξ ορισμού, αριθμητικές μεταβλητές στοιχίζονται δεξιά και αλφαριθμητικές αριστερά.
- ➔ “Measure”: επιλέγουμε “Scale” για συνεχείς μεταβλητές, “Ordinal” για μεταβλητές διάταξης και “Nominal” για κατηγορικές μεταβλητές.

Στο αρχείο δεδομένων του παραδείγματος, εισάγουμε τα κατάλληλα ονόματα και σχόλια μεταβλητών και επιλέγουμε να μην εμφανίζονται δεκαδικά ψηφία στις μεταβλητές. Οι τιμές των κατηγορικών μεταβλητών καταχωρούνται όπως φαίνεται στο επόμενο παράθυρο (επιλέγοντας “Add” κάθε φορά που εισάγεται μια τιμή και το αντίστοιχο όνομα αυτής):



Τέλος, δηλώνουμε ότι η μεταβλητή «φύλο» είναι κατηγορική. Το παράθυρο με τις μεταβλητές και το αντίστοιχο παράθυρο δεδομένων θα έχουν τη μορφή που φαίνεται στα επόμενα:

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	ύψος	Numeric	8	0	ύψος σε εκατοστά	None	None	10	Right	Scale
2	βάρος	Numeric	8	0	βάρος σε κιλά	None	None	10	Right	Scale
3	φύλο	Numeric	8	0	φύλο	{1, άνδρας}...	None	8	Right	Nominal
4	ταυτότητα	String	7	0	αριθμός ταυτότητας	None	None	10	Left	Nominal

	ύψος	βάρος	φύλο	ταυτότητα
1	60	163	γυναίκα	M111222
2	85	185	άνδρας	E333444
3	70	178	άνδρας	K444555
4	81	175	άνδρας	A222333
5	54	157	γυναίκα	M444222

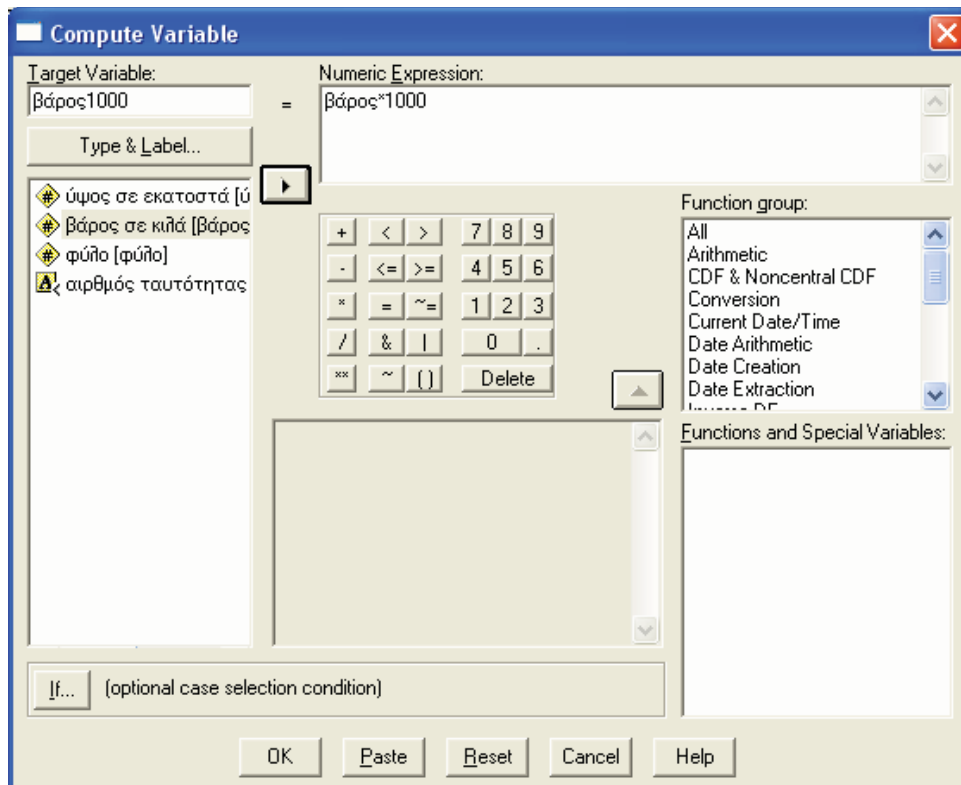
Τα ονόματα των τιμών της κατηγορικής μεταβλητής «φύλο» εμφανίζονται στο παράθυρο δεδομένων. Μπορούμε να επιλέξουμε να εμφανίζονται είτε οι τιμές είτε τα ονόματα αυτών από το μενού “View” και “Value labels”.

Μετά την ολοκλήρωση της εισαγωγής και επεξεργασίας των δεδομένων, το αρχείο που θα αποθηκεύσουμε θα πάρει την κατάληξη “.sav”, το οποίο υποδηλώνει αρχείο δεδομένων του SPSS. Υπάρχουν επίσης τα αρχεία:

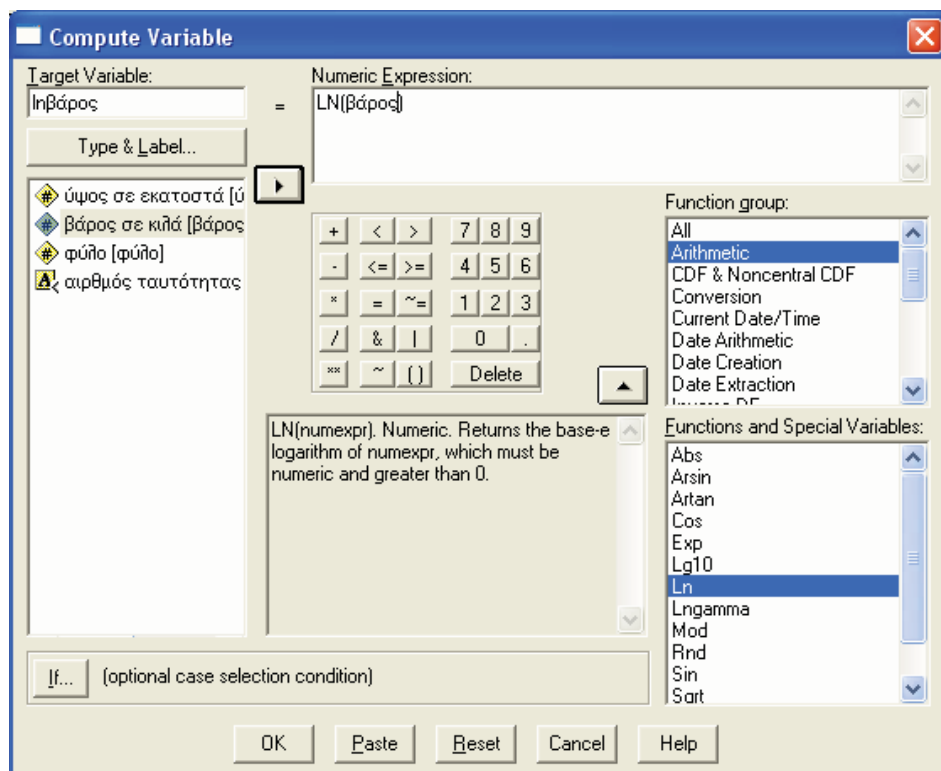
- ➔ “.spo” στα οποία αποθηκεύονται τα αποτελέσματα των αναλύσεων και
- ➔ “.sps” στα οποία αποθηκεύονται εντολές του SPSS.

Πράξεις μεταβλητών

Από το μενού “Transform” και “Compute”, παρέχεται η δυνατότητα να εκτελέσουμε πράξεις μεταξύ των υπάρχοντων μεταβλητών, είτε μετασχηματισμούς αυτών. Έστω για παράδειγμα ότι θέλουμε να δημιουργήσουμε τη μεταβλητή «βάρος1000», η οποία θα είναι το βάρος, εκφρασμένο σε γραμμάρια. Η εντολή που θα εισάγουμε είναι:

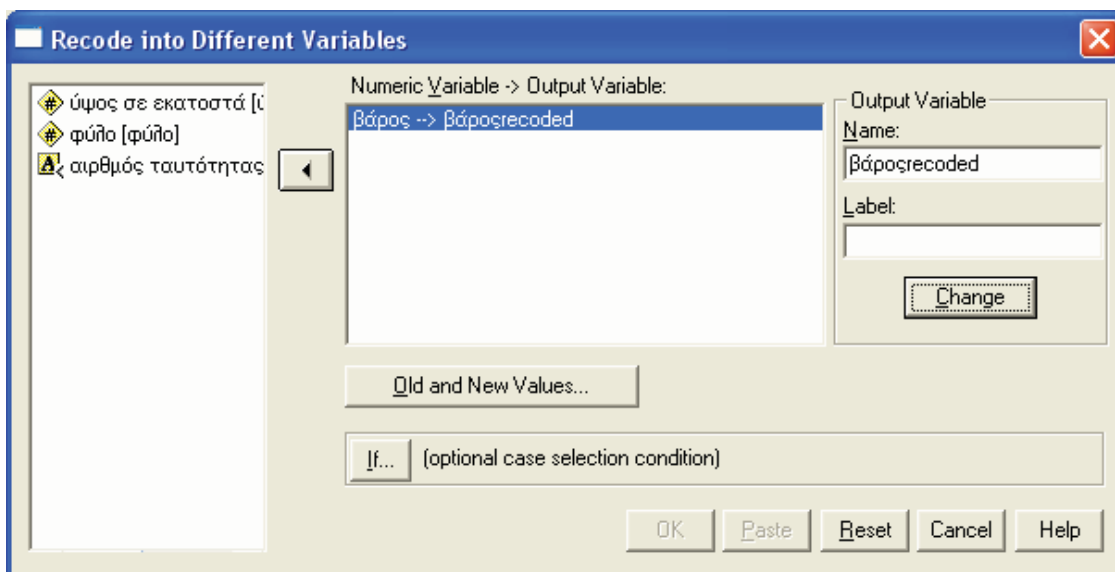


Επίσης, μπορούμε να μετασχηματίσουμε τις υπάρχουσες τιμές χρησιμοποιώντας τις συναρτήσεις που παρέχονται. Για παράδειγμα, ο νεπέριος λογάριθμος των τιμών του βάρους θα ληφθεί με την εντολή:

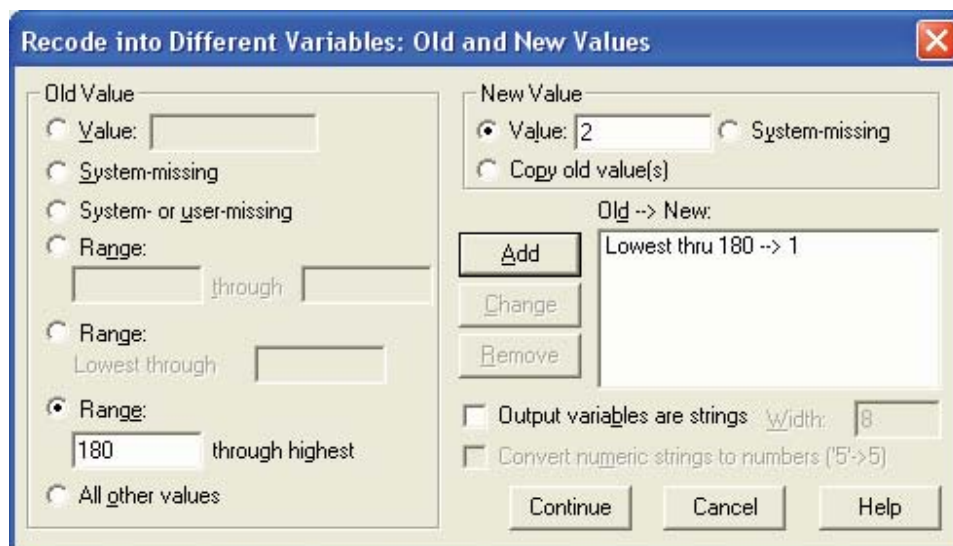


Κωδικοποίηση τιμών

Στο μενού “Transform” και “Recode” παρέχεται η δυνατότητα κωδικοποίησης των τιμών μιας μεταβλητής. Η νέα μεταβλητή μπορεί να αποθηκευτεί στη θέση της πρώτης μεταβλητής (αντικαθιστώντας την) είτε ως νέα (επιπρόσθετη) μεταβλητή. Αν για παράδειγμα θέλουμε να εντάξουμε σε δύο κατηγορίες τα άτομα του δείγματος ανάλογα με το αν έχουν ύψος λιγότερο ή περισσότερο από 180 εκατοστά, θα χρησιμοποιήσουμε τη διαδικασία που εμφανίζεται στην επόμενη εικόνα:



και στην επιλογή “Old and new values” θα δηλώσουμε:



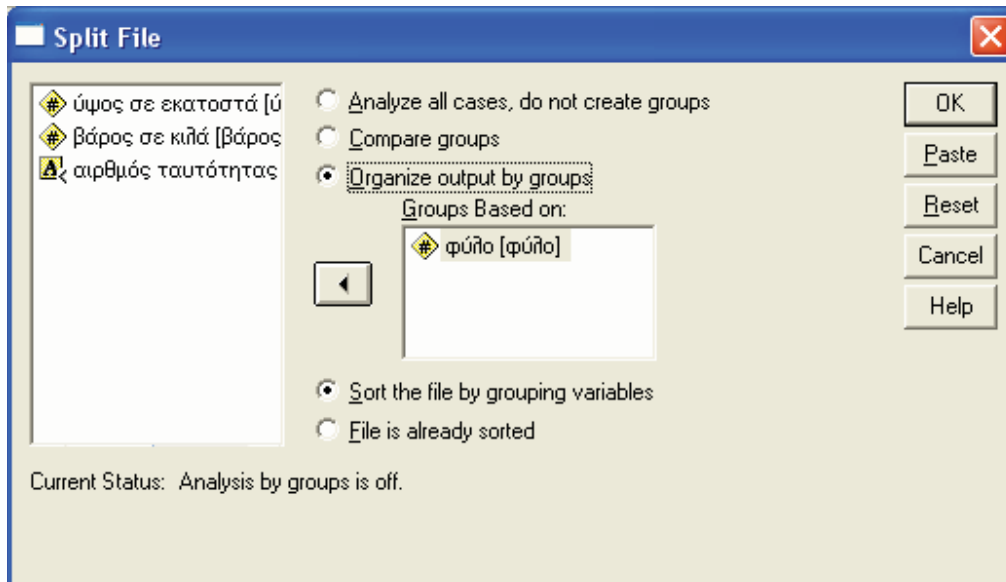
επιλέγοντας “Add” μετά από κάθε δήλωση νέας κατηγορίας. Η τιμή 180 εκατοστά που ορίζει την αλλαγή των διαστημάτων θα περιληφθεί (αν υπάρχει στο δείγμα) στο διάστημα στο οποίο χρησιμοποιείται ως δεξί άκρο. Οι δύο ομάδες δηλαδή που κατασκευάζονται με τον τρόπο αυτό έχουν κωδικούς:

- ➔ 1 αν το ύψος ανήκει στο διάστημα $(-\infty, 180]$
- ➔ 2 αν το ύψος ανήκει στο διάστημα $(180, +\infty)$.

Αναλύσεις για κατηγορίες ατόμων

Το SPSS παρέχει τη δυνατότητα αναλύσεων χωριστά για κατηγορίες ατόμων. Αν για παράδειγμα θέλουμε οι αναλύσεις που ζητάμε να γίνονται χωριστά για άνδρες και

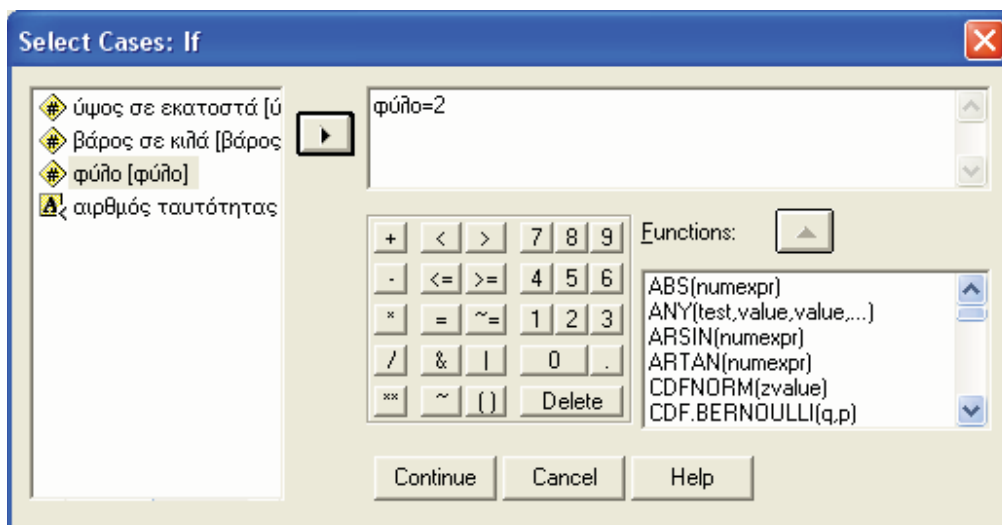
γυναίκες (και όχι σαν ενιαίο δείγμα), επιλέγουμε από το μενού “Data” και “Split file”. Στη συνέχεια, επιλέγοντας το “Organize output by groups”, δηλώνουμε τη μεταβλητή βάσει της οποίας θα γίνει ο διαχωρισμός του δείγματος σε υποομάδες:



Θα πρέπει να σημειωθεί ότι, μετά την επιλογή για χωριστές αναλύσεις ανά κατηγορία, ο χρήστης δεν θα παρατηρήσει καμία αλλαγή στο αρχείο των δεδομένων είτε των αποτελεσμάτων. Το αποτέλεσμα της διαδικασίας αυτής θα γίνει εμφανές αφότου ζητηθεί και ληφθεί συγκεκριμένη ανάλυση, η οποία θα εμφανίζεται για κάθε κατηγορία χωριστά.

Αναλύσεις σε επιλεγμένα άτομα

Στην περίπτωση που θέλουμε να συμπεριλάβουμε μόνο κάποια ή κάποιες ομάδες ατόμων σε αναλύσεις, επιλέγουμε από το μενού “Data” και “Select cases”. Επιλέγοντας το “If condition is satisfied”, δηλώνουμε τη συνθήκη που θα πρέπει να ικανοποιείται προκειμένου να περιληφθεί ένα άτομο στην ανάλυση. Αν για παράδειγμα θέλουμε να συμπεριληφθούν μόνο οι γυναίκες στην ανάλυση, το παράθυρο με την εντολή θα είναι:



Επίσης, στην εντολή επιλογής υποδείγματος, μπορούν να χρησιμοποιηθούν λογικοί τελεστές και τελεστές σύγκρισης. Αν, για παράδειγμα, έχουμε τέσσερις κατηγορίες σε μια μεταβλητή. Η εντολή:

- ➔ *(Μεταβλητή=1 OR μεταβλητή=3)* θα συμπεριλάβει στην ανάλυση μόνο τα άτομα που ανήκουν στις ομάδες 1 και 3
- ➔ *(Μεταβλητή>=3)* ή ισοδύναμα *(Μεταβλητή=3 OR μεταβλητή=4)* θα συμπεριλάβει στην ανάλυση μόνο τα άτομα που ανήκουν στις ομάδες 3 και 4
- ➔ *Μεταβλητή=2 AND φύλο=2* θα συμπεριλάβει στην ανάλυση τις γυναίκες που ανήκουν στη δεύτερη ομάδα
- ➔ *Μεταβλητή=1 AND μεταβλητή=3* θα συμπεριλάβει στην ανάλυση τα άτομα που ανήκουν συγχρόνως και στην ομάδα 1 και στην ομάδα 3 (δηλαδή δεν θα περιληφθεί κανένα άτομο).

Εναλλακτικά, αντί των τελεστών AND και OR, μπορούν να χρησιμοποιηθούν τα σύμβολα & και | αντίστοιχα.

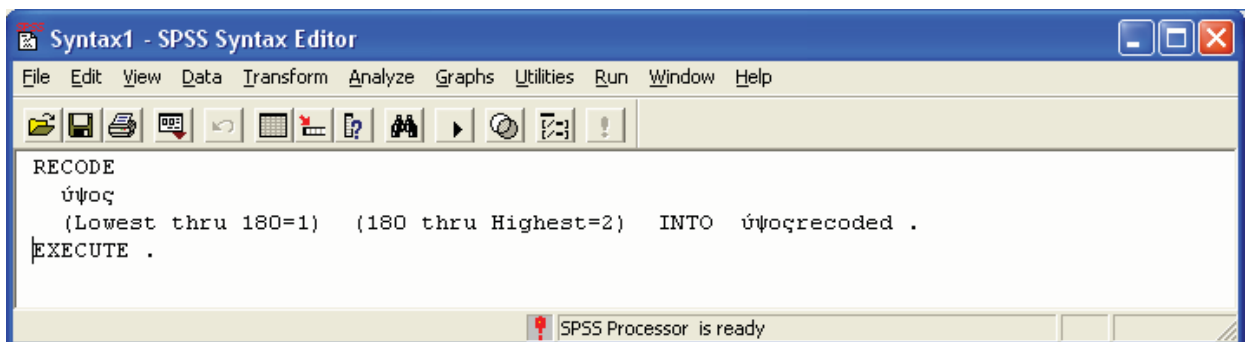
Μετά την εφαρμογή της διαδικασίας αυτής, οι γραμμές των ατόμων που δεν επιλέγονται για να χρησιμοποιηθούν, εμφανίζονται με μια γραμμή διαγραφής στον αύξοντα αριθμό τους. Αν, για παράδειγμα, επιλέξουμε μόνο τις γυναίκες, το αρχείο δεδομένων θα έχει τη μορφή:

	ύψος	βάρος	φύλο	ταυτότητα	filter_\$
1	60	163	γυναίκα	M111222	Selected
2	85	185	άνδρας	E333444	Not Selecte
3	70	178	άνδρας	K444555	Not Selecte
4	81	175	άνδρας	A222333	Not Selecte
5	54	157	γυναίκα	M444222	Selected

Η μεταβλητή “filter_\$” έχει δημιουργηθεί αυτόματα από το στατιστικό πακέτο, η οποία υποδηλώνει αν το καθένα άτομο περιλαμβάνεται ή όχι στην ανάλυση.

Αρχεία εντολών του SPSS

Σε κάθε διαδικασία χειρισμού μεταβλητών είτε στατιστικών αναλύσεων, το SPSS παρέχει τη δυνατότητα να χρησιμοποιήσουμε εντολές. Αν αντί του πλήκτρου “OK”, έχοντας ζητήσει μια διαδικασία, επιλέξουμε το πλήκτρο “Paste”, το πακέτο ανοίγει ένα παράθυρο με την εντολή. Για παράδειγμα, στη δημιουργία δύο ομάδων ατόμων βάσει του ύψους τους, αν επιλεγθεί το πλήκτρο “Paste”, θα εμφανιστεί το παράθυρο των εντολών:



```

Syntax1 - SPSS Syntax Editor
File Edit View Data Transform Analyze Graphs Utilities Run Window Help
[Icons]
RECODE
  ύψος
  (Lowest thru 180=1) (180 thru Highest=2) INTO ύψοςrecoded .
EXECUTE .
SPSS Processor is ready

```

Επιλέγοντας την εντολή και στη συνέχεια “Run” από το μενού, θα εκτελεσθεί η εντολή. Η διαδικασία αυτή είναι εξαιρετικά χρήσιμη όταν χρειάζεται να πραγματοποιήσουμε πολλές φορές την ίδια διαδικασία σε διάφορες μεταβλητές. Με χρήση εντολών, μπορούμε να αντιγράψουμε την εντολή όσες φορές χρειάζεται και απλά να αλλάζουμε στην κάθε αντιγραμμένη εντολή, το όνομα της μεταβλητής κατάλληλα.

Τέλος Ενότητας



Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο



Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

